

PH <i>US</i> 0 10057W0	MAT. DOSSIER
---------------------------	-----------------

①9 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

①2 Offenlegungsschrift
①1 DE 37 27 530 A1

①5 Int. Cl. 4:
H04N 7/13
// H04N 11/04

②1 Aktenzeichen: P 37 27 530.5
②2 Anmeldetag: 18. 8. 87
④3 Offenlegungstag: 2. 3. 89

Patentamt

DE 37 27 530 A1

⑦1 Anmelder:
Philips Patentverwaltung GmbH, 2000 Hamburg, DE

⑦2 Erfinder:
Vogel, Peter, Dipl.-Ing., 8566 Diepersdorf, DE

⑤4 Verfahren zur Bestimmung von Bewegungsvektoren

Das beschriebene Verfahren zur Bestimmung von Bewegungsvektoren wird bei der blockweisen Codierung von Videobildern angewendet. Ein dem Codierer zugeführter Eingangsblock wird mit gleichgroßen Bildausschnitten des vorangegangenen Videobildes verglichen und dem Eingangsblock derjenige Vektor als Bewegungsvektor zugeordnet, der die Verschiebung von der Position des Eingangsblockes innerhalb eines Videobildes zu demjenigen Bildausschnitt angibt, der die größte Ähnlichkeit mit dem Eingangsblock hat.

Um die Bewegungsvektoren für sämtliche Eingangsblocke eines Videobildes in nur einem Durchgang zu bestimmen und dabei ein glattes Feld von Bewegungsvektoren zu erhalten, wird jedem Eingangsblock derjenige Vektor v als Bewegungsvektor zugeordnet, der die Summe aus dem Prädiktionsfehler zwischen dem Eingangsblock und einem gesuchten Bildausschnitt sowie den gewichteten Abweichungen des Vektors v von schon bestimmten Bewegungsvektoren zum Minimum macht. Der gesuchte Bildausschnitt ist dabei derjenige Ausschnitt, der durch Verschiebung um den Vektor v aus der Position des Eingangsblocks hervorgeht.

DE 37 27 530 A1

Patentansprüche

1. Verfahren zur Bestimmung von Bewegungsvektoren bei der blockweisen Codierung von Videobildern, bei dem ein Eingangsblock b_i mit gleichgroßen Bildausschnitten des vorangegangenen Videobildes verglichen wird, und dem Eingangsblock b_i derjenige Vektor als Bewegungsvektor v_i zugeordnet wird, der die Verschiebung von der Position i zu demjenigen Bildausschnitt angibt, der die größte Ähnlichkeit mit dem Eingangsblock b_i hat, dadurch gekennzeichnet, daß unter der Ähnlichkeit zwischen dem Eingangsblock b_i und einem gleichgroßen Bildausschnitt c der Ausdruck

$$E = \|b_i - c\| + \sum_{k \neq i} f_k \|v - v_k\|$$

verstanden wird, wobei der erste Summand den Prädiktionsfehler angibt und der zweite Summand die mit Koeffizienten f_k gewichtete Summe der Abweichungen eines Vektors v von schon bestimmten Bewegungsvektoren v_k ($k \neq i$) anderer Eingangsblöcke darstellt, daß der Vektor v die Verschiebung von der Position i zur Position des Bildausschnittes c angibt und daß dem Eingangsblock b_i derjenige Vektor v als Bewegungsvektor v_i zugeordnet wird, der den Ausdruck E zum Minimum macht.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Summe im Ausdruck E zwei Terme enthält und es sich bei den schon bestimmten Bewegungsvektoren v_i und v_m um die Bewegungsvektoren derjenigen Blöcke handelt, die der Position i am nächsten liegen.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß die in der Summe des Ausdrucks E verbleibenden beiden Koeffizienten f_i und f_m derart von den Bewegungsvektoren v_i und v_m abhängig gemacht werden, daß ein Koeffizient den Wert Null bekommt, wenn der zugehörige Bewegungsvektor der Nullvektor ist und einen von Null verschiedenen Wert bekommt, wenn der zugehörige Bewegungsvektor nicht der Nullvektor ist.

4. Verfahren nach Anspruch 3, dadurch gekennzeichnet, daß die Bewegungsvektoren v_i der Eingangsblöcke b_i in der Reihenfolge bestimmt werden, daß — angefangen bei einem Randblock — alle Blöcke einer Blockzeile nacheinander in einer Richtung durchlaufen werden und daß die nächste Blockzeile in umgekehrter Richtung durchlaufen wird.

Beschreibung

Die Erfindung betrifft ein Verfahren zur Bestimmung von Bewegungsvektoren bei der blockweisen Codierung von Videobildern, bei dem ein Eingangsblock b_i mit gleichgroßen Bildausschnitten des vorangegangenen Videobildes verglichen wird, und dem Eingangsblock b_i derjenige Vektor als Bewegungsvektor v_i zugeordnet wird, der die Verschiebung von der Position i zu demjenigen Bildausschnitt angibt, der die größte Ähnlichkeit mit dem Eingangsblock b_i hat.

Ein solches Verfahren ist Ausgangsbasis einer Reihe von Verbesserungsvorschlägen bei der Codierung von Videobildern. Das durch die angegebenen Merkmale

charakterisierte Verfahren hat — wie weiter unten näher erläutert werden wird — unter anderem der Nachteil, daß besonders bei niedrigen Übertragungsbitraten (z. B. bei 64 kBit/s) die Bildqualität sehr mangelhaft sein kann, weil bei der Wiedergabe von bewegten Objekten auf dem Bildschirm z. B. Blockränder störend sichtbar werden.

Um derartige Störungen abzuschwächen, hat W. Geuen (Geuen, W.: Postprocessing of Motion Vectors. Poster des Forschungsinstitutes der Deutschen Bundespost, P.O. Box 5000, DE-6100 Darmstadt) vorgeschlagen, daß auch der Gesamtheit aller Bewegungsvektoren eines Videobildes bestehende Vektorfeld nachträglich zu glätten. Dabei wird von der Vorstellung ausgegangen, daß sich das Vektorfeld — wenn es der tatsächlichen Bewegung entspricht — von Block zu Block nur wenig ändert, sofern die Blöcke benachbart sind. Ausreißer — also wesentliche Abweichungen eines Bewegungsvektors von den Bewegungsvektoren der Nachbarblöcke — kommen nicht vor. Daher wird mit speziellen Filtern nachträglich das abgespeicherte Vektorfeld geglättet, d. h., nach ihrer üblichen Bestimmung werden die Bewegungsvektoren durch ein spezielles Verfahren so abgeändert, daß Bewegungsvektoren benachbarter Blöcke nur gering voneinander abweichen. In der Tat läßt sich die Bildqualität damit verbessern. Zu beachten ist, daß mit der nachträglichen Änderung der Bewegungsvektoren der Prädiktionsfehler (siehe weiter unten) nicht immer minimal ist.

Einen anderen Weg schlagen Brandt et al. ein (v. Brandt A., und Templer, W.: Obtaining Smoothed Optical Flow Fields by Modified Block Matching. 5th SCIA, Stockholm (1987) Juni 2—5, Seiten 523—529). Von diesen Verfassern wird vorgeschlagen, beim Aufsuchen des Bewegungsvektors eines Blocks von vornherein anders vorzugehen. Während bisher die L1-Norm (näheres siehe weiter unten) von Differenzblöcken minimiert wurde, wird zu dieser Norm nun noch ein gesonderter Term addiert, der — zum Zwecke der Glättung — von den Komponenten der Bewegungsvektoren vier benachbarter Blöcke abhängt. Als vorläufiger Prädiktionsblock mit vorläufigem Bewegungsvektor wird derjenige Block angesehen, der den Gesamtausdruck (L1-Norm erweitert um den gesonderten Term) zum Minimum macht. Das Ergebnis wird — wie schon angedeutet — jedoch nicht als endgültig angesehen, sondern in theoretisch unendlich vielen Durchgängen für das gesamte Bild wiederholt, bis sich keine Änderungen mehr ergeben. Sodann wird die Blockgröße verringert und das Verfahren unter Verwendung des bisherigen Ergebnisses mit der verkleinerten Blockgröße wiederholt.

Der Erfindung liegt die Aufgabe zugrunde, ein Verfahren der eingangs genannten Art anzugeben, bei dem ein geglättetes Feld von Bewegungsvektoren bei nur einmaligem Durchgang aller Blöcke eines Videobildes erhalten wird.

Diese Aufgabe wird dadurch gelöst, daß unter der Ähnlichkeit zwischen dem Eingangsblock b_i und einem gleichgroßen Bildausschnitt c der Ausdruck

$$E = \|b_i - c\| + \sum_{k \neq i} f_k \|v - v_k\|$$

verstanden wird, wobei der erste Summand den Prädiktionsfehler angibt und der zweite Summand die mit Koeffizienten f_k gewichtete Summe der Abweichungen eines Vektors v von schon bestimmten Bewegungsvektoren

ren v_k ($k=i$) anderer Eingangsblöcke darstellt, daß der Vektor v die Verschiebung von der Position i zur Position des Bildausschnittes c angibt und daß dem Eingangsblock b_i derjenige Vektor v als Bewegungsvektor v_i zugeordnet wird, der den Ausdruck E zum Minimum macht.

Vorteilhafte Ausgestaltungen des erfindungsgemäßen Verfahrens sind in den Unteransprüchen angegeben. Anhand der Figur und eines Ausführungsbeispiels soll die Erfindung näher erläutert werden.

Die Figur zeigt einen Hybrid-Codierer, bei dem das erfindungsgemäße Verfahren zur Anwendung kommt.

Der Hauptzweck des in der Figur gezeigten Hybrid-codierers ist der, die von einer Video-Datenquelle kommenden Videodaten möglichst mit geringem Informationsverlust in ein Signal mit möglichst geringer Bitrate umzu codieren (vergleiche z. B. die deutsche Patentanmeldung DE 36 13 344). Bei diesem Vorgang werden zwei Codierungsprinzipien — daher der Name Hybrid-Codierer — angewendet:

Das Interframe-Prinzip, bei dem die Korrelation zwischen zeitlich aufeinanderfolgenden Videobildern (diese Bezeichnung wird hier für Voll- und Teilbilder verwendet) ausgenutzt wird und das Intraframe-Prinzip, bei dem die Korrelation der Videodaten innerhalb eines Videobildes ausgenutzt wird.

Vor dem eigentlichen Codierungsprozeß ist eine Aufbereitung der Daten erforderlich. Diese Aufbereitung wird bei dem Hybrid-Codierer nach der Figur durch eine Funktionseinheit PP (preprocessing) übernommen. Die Daten werden in Blöcken an den Codierer übergeben. Ein jeder dieser Videodatenblöcke enthält die Daten bestimmter Bildpunkte eines Videobildes, die als Elemente einer quadratischen Zahlenmatrix aufgefaßt werden und die einen in seiner Position festgelegten Ausschnitt aus einem Videobild darstellen (zur Bedeutung der hier verwendeten Begriffe im Zusammenhang mit Zahlenmatrizen vgl. Wigner, E.P.: Group Theory; Academic Press New York and London 1959, S. 1—30).

So kann z. B. ein Datenblock aus den Chrominanzwerten der ersten acht Bildpunkte der ersten acht Zeilen eines Videobildes bestehen. Durch die Funktionseinheit PP wird jedes Videobild in gleich große Datenblöcke, d. h., in gleichgroße Ausschnitte mit vorbestimmter Position, zerlegt. Die Zerlegungsvorschrift läßt auch die Position — also z. B. die Mittelpunktskoordinaten — eines Datenblocks innerhalb eines gesamten Videobildes erkennen. Der eben als Beispiel angeführte Datenblock könnte z. B. durch b_1 symbolisiert werden, wobei der Index 1 als Kurzzeichen für die Mittelpunktskoordinaten des Datenblocks aufgefaßt werden kann. Datenblöcke von aufeinanderfolgenden Videobildern, die den gleichen Index haben, werden hier als sich entsprechende Datenblöcke bezeichnet.

Mit dergleichen Bezeichnung sollen auch Datenblöcke aufeinanderfolgender Videobilder gemeint sein, deren Informationsinhalte die größte Übereinstimmung haben, deren Indices jedoch nicht übereinzustimmen brauchen. In diesem Sinne sich entsprechende Datenblöcke spielen bei Hybrid-Codierern eine Rolle, bei denen ein sogenanntes blockmatching vorgenommen wird. Für diese Variante von Hybrid-Codierern ist das vorliegende Verfahren gedacht.

Bei der Übergabe eines Datenblockes z. B. des Datenblocks b_1 an einen Eingang eines Subtrahierers SR wird gleichzeitig der entsprechende Datenblock — symbolisiert z. B. durch b_1 — des vorangegangenen Videobildes aus einem Bildspeicher BS an den anderen Eingang des

Subtrahierers SR gegeben. In diesem Beispiel hat also der entsprechende Datenblock b_1 den gleichen Index — also die gleiche Position — wie der Eingangsblock b_1 . Der Subtrahierer SR bildet die Differenz zwischen den beiden Blöcken im Sinne der Differenz zwischen zwei Matrizen (vgl. Wigner l.c., S. 7); dieser Differenzblock wird nun weiteren Operationen unterworfen.

Eine erste Operationseinheit $OE1$ nimmt mit jedem Differenzblock eine Ähnlichkeitstransformation im Sinne einer Matrixtransformation vor (vgl. Wigner l.c., S. 9). Sei a das Symbol für die Transformationsmatrix der Einheit $OE1$ und d_1 das Symbol für die Matrix des Differenzblockes — oder einfacher für den Differenzblock — so liegt nach der Transformation am Ausgang der Einheit $OE1$ der Block $D_1 = a^{-1} d_1 a$ an, wobei a^{-1} das Symbol für die zu a inverse Matrix bedeutet. Die Transformation durch die Einheit $OE1$ entspricht etwa der Fourier-Transformation bei der akustischen Signalübertragung; der Block D_1 kann meist mit weniger Binärstellen codiert werden als der Block d_1 .

Anschließend durchläuft das transformierte Signal einen Quantisierer Q , der nochmals für eine Datenreduktion sorgt. Damit das gesamte Signal mit konstanter Bitrate an einen Empfänger übertragen werden kann, ist ein Pufferspeicher P vorgesehen. Ein Multiplexer MUX verschachtelt das aus dem Pufferspeicher P ausgelesene Nutzsignal mit Steuerinformationen, die im vorliegenden Falle unter anderem zur Einstellung eines Quantisierers auf der Empfängerseite dienen (es handelt sich hier um adaptive Quantisierer). Nach der Quantisierung wird das Signal auch über einen Rückkopplungsweg an den Eingang des Hybrid-Codierers zurückgeschleift. Zunächst wird der durch den Quantisierer Q veränderte Block D_1 durch eine nicht eingeleitete Regenerationseinheit so weit regeneriert, daß er bis auf Rundungsfehler mit dem ursprünglichen Block D_1 übereinstimmt. Sodann wird er durch eine zweite Operationseinheit $OE2$ mit der Transformationsmatrix a^{-1} wieder in den Differenzblock d_1 (ebenfalls bis auf Rundungsfehler) zurücktransformiert. Ein Addierer AR addiert zu diesem Block — wegen der Verbindung eines Ausgangs A des Bildspeichers BS mit einem Eingang des Addierers AR — den Datenblock b_1 , mit dem durch den Subtrahierer SR der Differenzblock d_1 gebildet wurde. Eventuelle Verzögerungen wegen endlicher Laufzeiten werden entweder durch Verzögerungsglieder oder Taktverschiebungen (beides in Fig. 1 nicht angedeutet) ausgeglichen.

Wie sich leicht überprüfen läßt, ergibt sich am Ausgang des Addierers AR der Datenblock b_1 (bis auf Rundungsfehler) des gerade über die Funktionseinheit PP zugeführten Videobildes. Dieser Datenblock wird über einen Eingang E des Bildspeichers BS in ihn eingelesen und übernimmt dort die Rolle des Datenblockes b_1 , der nun gelöscht wird.

Die Vorrichtung, die dem Aufsuchen des entsprechenden Datenblocks — auch Prädiktionsblock genannt — dient, ist Bestandteil des Bildspeichers BS . Dem Bildspeicher BS wird nämlich über eine gesonderte Zuleitung — angedeutet durch die Verbindung des Ausgangs der Funktionseinheit PP mit dem Bildspeicher BS — ebenfalls der gerade erzeugte Eingangsblock b_i mit der Position i innerhalb eines Videobildes zugeführt. Ausgehend von der Position i werden nun im Bildspeicher BS nacheinander alle diejenigen Bildausschnitte von der Größe eines Eingangsblocks mit dem Eingangsblock b_i verglichen, die sich aus der Position i durch Antragen eines der Vektoren eines vorgegebenen Satzes von 225

Verschiebungsvektoren $g_1, g_2 \dots g_{225}$ ergeben. Selbstverständlich enthält dieser Satz von Vektoren auch den Nullvektor.

Die sich durch die Verschiebung mit den Vektoren $g_1, g_2 \dots g_{225}$ ergebenden Bildausschnitte stimmen in ihrer Lage im allgemeinen nicht mit der Lage eines Blockes überein, vielmehr sind diese Bildausschnitte gegen das Blockraster versetzt. Die minimale von Null verschiedene Versetzung ist so groß wie der Abstand zwischen zwei Bildpunkten.

Der Bildausschnitt des gespeicherten Bildes, der sich durch Verschiebung aus der Position i um den Vektor \underline{V} ergibt, soll mit $c_i(\underline{V})$ bezeichnet werden. Den Eingangsblock b_i mit dem Block $c_i(\underline{V})$ vergleichen heißt zunächst, die L1-Norm des Differenzblockes $u_i = b_i - c_i(\underline{V})$ zu bilden. Darunter wird die Summe der Beträge aller Elemente des Differenzblockes verstanden, die durch die symbolische Schreibweise u_{iL1} angedeutet wird. Soll die Art der Norm offenbleiben, wird der Index L1 fortgelassen. Die Norm des oben angegebenen Differenzblockes wird auch Prädiktionsfehler genannt, weil der Block $c_i(\underline{V})$ beim Codieren als Vorhersageblock (Prädiktionsblock) für den Block b_i verwendet werden kann. Zu diesem Prädiktionsfehler werden noch zwei weitere Terme addiert, die vom gewichteten Unterschied des Vektors \underline{V} zwischen zwei schon bekannten Vektoren \underline{v}_j und \underline{v}_m abhängen. Insgesamt ergibt sich nun der Ausdruck

$$E = \| b_i - c_i(\underline{V}) \| + f_j \cdot \| \underline{V} - \underline{v}_j \| + f_m \cdot \| \underline{V} - \underline{v}_m \|, \quad 30$$

wobei hier das Symbol $\| \dots \|$ wiederum die Betragssumme der Komponenten des Vektors bedeutet, der an der Stelle der Punkte steht. Die Koeffizienten f_j und f_m sind die Gewichte, mit denen die Norm der Differenzvektoren gewichtet wird. Als Bewegungsvektor \underline{v}_j wird nun derjenige Verschiebungsvektor $g_1, g_2 \dots g_{225}$ dem Eingangsblock b_i zugeordnet, der — an der Stelle von \underline{V} eingesetzt — den Ausdruck E zum Minimum macht. Die Vektoren \underline{v}_j und \underline{v}_m sind die Bewegungsvektoren jener schon bearbeiteter Eingangsböcke, die der Position i am nächsten liegen. Da zunächst unterstellt wird, daß die Blöcke eines Videobildes von links nach rechts sowie von oben nach unten abgearbeitet werden, sind diese Vektoren den Blöcken zugeordnet, von denen einer über dem Block b_i liegt und der andere links daneben. 40

Ist der Block b_i Randblock, wird mindestens einer der Vektoren \underline{v}_j oder \underline{v}_m durch den Nullvektor ersetzt.

Werden die Bewegungsvektoren aller Blöcke durch Minimierung des Ausdrucks E mit festen Koeffizienten f_j und f_m gewonnen, so ergibt sich ein glattes Vektorfeld, das mit nur einem Durchgang durch alle Blöcke des Videobildes erhalten wird. Dennoch erhält man bei der Bildwiedergabe Artefakte (nicht näher angebbare Bildstörungen) an den Übergangszonen zwischen bewegten und unbewegten Objekten. In diesem Bereich sind aufgrund der Veränderungen von Bild zu Bild große Bewegungsvektoren zu erwarten. Im Gegensatz dazu liefert der Ausdruck E mit festen Koeffizienten f_j und f_m zu oft Nullvektoren als Bewegungsvektoren. Zur Behebung dieses Nachteils ist vorgesehen, Koeffizienten f_j und f_m in folgender Weise von den schon bestimmten Bewegungsvektoren \underline{v}_j und \underline{v}_m abhängig zu machen: 60

$$f_i = \begin{cases} 0 & \text{für } \underline{v}_i = (0, 0) \\ f & \text{sonst} \end{cases} \quad 65$$

Entsprechendes gilt für f_m .

Mit dieser Kopplung der Konstanten f_j und f_m an die Vektoren \underline{v}_j und \underline{v}_m der schon abgearbeiteten Blöcke ergeben sich an den Rändern bewegter Objekte zwar bessere Konturen als vorher, jedoch entstehen gerade an den Rändern Bewegungsvektoren, die nicht annähernd mit der Richtung der Bewegungsvektoren im Innern der bewegten Objekte übereinstimmen. Diese Ausreißer von Vektoren treten nur dann auf, wenn man sich bei der Abarbeitung eines Gesamtbildes von links dem bewegten Objekt nähert, also nur am linken Rand des bewegten Objektes. Ändert man die Richtung, in der die Blöcke abgearbeitet werden, so treten die Ausreißer nur am rechten Rand auf. Die beschriebenen Ausreißer von Vektoren führen zwar zu weniger starken Artefakten an den Rändern bewegter Objekte als Nullvektoren, jedoch läßt sich auch der Rest an Artefakten noch beseitigen, wenn folgende Reihenfolge der Abarbeitung eingehalten wird:

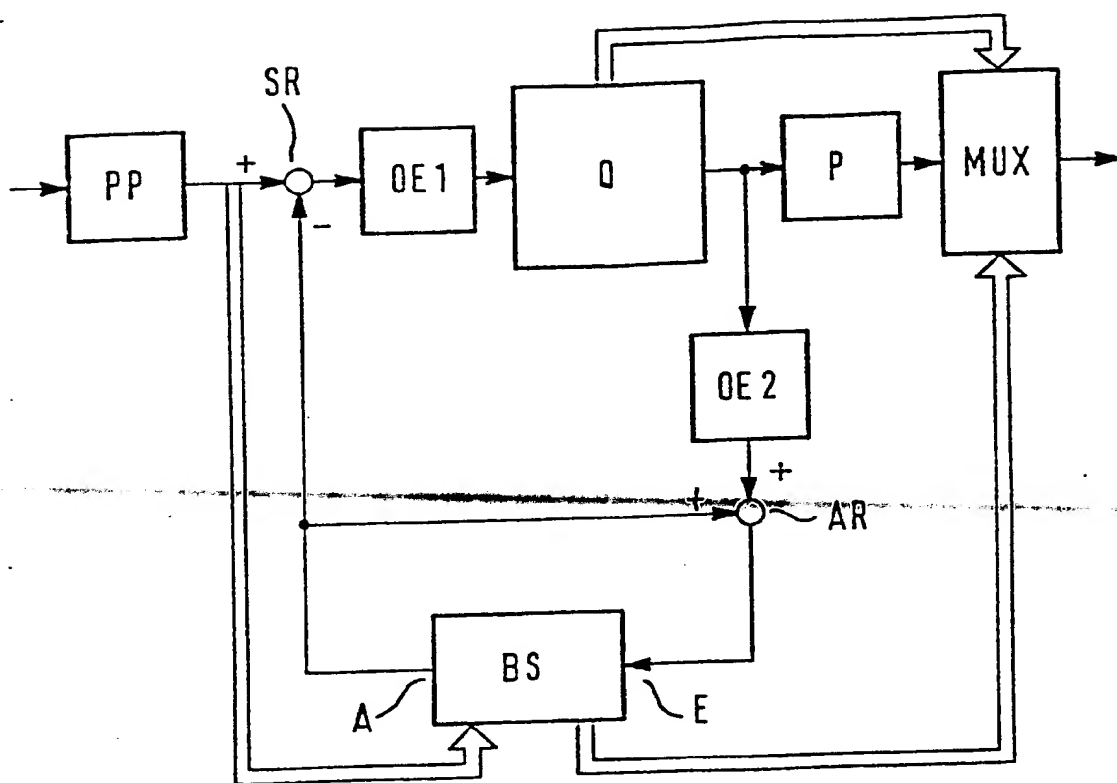
Man beginnt mit der ersten Blockreihe von links nach rechts in der obersten Reihe des Videobildes und arbeitet die zweite Blockreihe in der entgegengesetzten Richtung durch. Die Abarbeitungsrichtung der dritten Reihe stimmt wieder mit der ersten Reihe überein. Es wird also die Richtung der Abarbeitung geändert, sobald man dabei an einen Rand des Bildes gerät.

- Leerseite -

Nummer:
Int. Cl.4:
Anmeldetag:
Offenlegungstag:

37 27 530
H 04 N 7/13
18. August 1987
2. März 1989

3727530



PHD 87 160